

### ***Supplementary materials***

**Supplementary Table 1.** Distribution of TFs in sequenced prokaryotic genomes. Total number of ORFs and number of predicted TFs per genome are from DBD database.<sup>56</sup> Genomes are sorted according to taxonomic groups. Obligate pathogens and endosymbionts with genome size less than 1500 ORFs are marked in grey background. Pathogenic phenotype, ecotype and biotechnological and medical relevance are from the Integrated Microbial Genomes web-site of Joint Genomic Institute (<http://img.jgi.doe.gov/>).

**Supplementary Table 2.** Distribution of sigma factors in sequenced bacterial genomes. Sigma factors from three different families ( $\sigma^{70}$ ,  $\sigma^{54}$ , and ECF) were found in bacterial genomes using HMM profiles based on experimentally verified sigma factors<sup>77</sup> and deposited in the Genome Atlas Database (<http://www.cbs.dtu.dk/services/GenomeAtlas/>).<sup>286</sup> Obligate pathogens and endosymbionts with genome size less than 1500 ORFs are marked in grey background.

Dmitry A. Rodionov was born in Moscow, Russia in 1977. He studied biophysics in Moscow Engineering Physics Institute and received his M.Sc. degree in 2000, analyzing transcriptional regulation of multidrug resistance genes in *Escherichia coli* under the supervision of Dr.

Alexandra B. Rakhmaninova. His graduate research was conducted with Prof. Mikhail S. Gelfand and Prof. Andrei A. Mironov in Laboratory of Bioinformatics of the State Research Institute for Genetics and Selection of Industrial Microorganisms in Moscow, Russia. At the same institution, he obtained his Ph.D. in Molecular Biology in 2003 with a thesis about comparative genomics reconstruction of transcriptional regulatory and metabolic networks in bacteria. He continued to study microbial regulatory networks using comparative genomics with Prof. Gelfand at the Institute for Information Transmission Problems of Russian Academy of Sciences in Moscow. During this time, he participated in multiple collaborative projects as a visitor scientist in National Center for Biotechnology Information in Bethesda, Maryland (spring 2000), Unit of Microbiology and Genetics, INSA-Lyon, France (autumn 2003), Lawrence Berkeley National Laboratory in California (spring 2004), and Institute for Microbiology at Humboldt University in Berlin, Germany (spring 2005). In autumn 2005, he moved to the laboratory of Prof. Andrei Osterman in the Burnham Institute for Biomedical Research in La Jolla, California as a postdoctoral associate. His research interests are different aspects of comparative genomic reconstruction of metabolic pathways and regulatory networks in prokaryotes including identification of transcription factor-binding sites and RNA regulatory elements, functional annotation of genes and identification of missing genes in metabolic pathways, phylogenetic analysis of protein families, and evolution of metabolic pathways and regulatory systems.